

# Accurate Sampling Timing Acquisition for Baseband OFDM Power-line Communication in Non-Gaussian Noise

Chen Chen, Yun Chen, *Member, IEEE*, Na Ding, Yizhi Wang, Jia-Chin Lin, *Senior Member, IEEE*, Xiaoyang Zeng, *Member, IEEE*, and Defeng (David) Huang, *Senior Member, IEEE*

**Abstract**—In this paper, a novel technique is proposed to address the joint sampling timing acquisition for baseband and broadband power-line communication (BB-PLC) systems using Orthogonal-Frequency-Division-Multiplexing (OFDM), including the sampling phase offset (SPO) and the sampling clock offset (SCO). Under pairwise correlation and joint Gaussian assumption of received signals in frequency domain, an approximated form of the log-likelihood function is derived. Instead of a high complexity two-dimension grid-search on the likelihood function, a five-step method is employed for accurate estimations. Several variants are presented in the same framework with different complexities. Unlike conventional pilot-assisted schemes using the extra phase rotations within one OFDM block, the proposed technique turns to the phase rotations between adjacent OFDM blocks. Analytical expressions of the variances and biases are derived. Extensive simulation results indicate significant performance improvements over conventional schemes. Additionally, effects of several noise models including non-Gaussianity, cyclo-stationarity, and temporal correlation are analyzed and simulated. Robustness of the proposed technique against violation of the joint Gaussian assumption is also verified by simulations.

**Index Terms**—OFDM, power-line communication, baseband system, sampling phase offset, sampling clock offset, non-Gaussian noise

## I. INTRODUCTION

Power-line communication (PLC) is a potential solution for long haul transmission, last mile access, and in-building connection over low to high voltage power networks. By turning virtually each wired device into a target of future value-added services, PLC acts as a technological enabler spanning internet access, residential or business premises, smart grid, and other municipal applications [1]. Moreover, it enjoys the advantages of pervasive power cable infrastructures and low implementation cost without rewiring [2].

To provide a higher data rate than earlier ultra narrow-band PLC (UNB-PLC) and narrowband PLC (NB-PLC) solutions, broadband PLC (BB-PLC) empowered by multi-carrier schemes increases the data rate to 200 Megabits per second (Mbps) and the frequency band to 1.8 – 30 MHz [3]. One step

further, the new ITU and IEEE standards extend the frequency band to 100 MHz and increase the data rate to 500 Mbps [4].

Originally designed for electricity delivery, power line is a harsh and noisy medium, particularly for low power and high frequency data transmission [1]. Statistical modeling of PLC channels is a technical challenge for communication theory approach [4]. In [5], Zimmermann et al. suggested an analytical model for the channel transfer function (CTF) of PLC, characterized by a small number of parameters for frequency below 30 MHz. To broaden its coverage, Tonello et al. in [6] proposed a random broadband channel generator based on measurement campaign results on nine classes of channels in the 2 ~ 100 MHz range. [7], [8] discussed the time-variation of PLC channels. However, for broadband transmission, it could be largely neglected, since its period (reciprocal to the mains frequency) is much longer than the block duration.

Different types of noise are observable in power lines. A previous work [2] suggested the Nakagami-m distribution to fit the envelope probability density function (PDF) of time domain background noise. Its accuracy was proved in [9] by empirical measurements in a 10-kV medium-voltage power network. Relevant works can be found in [10], [11]. The Class-A impulsive noise model was proposed in [12]. It has been extensively used ever since due to its canonical and closed-form PDF [13], as well as excellent agreement with measurements from both natural and man-made noise environments [14]. The power spectral density (PSD) of impulsive noise can reach a value of more than 50 dB above background noise [15]. In [16], Katayama et al. revealed the cyclo-stationarity of the noise, synchronous to the mains voltage frequency. Also, noise samples are temporally correlated, as shown in [17].

Orthogonal-Frequency-Division-Multiplexing (OFDM) is one option for multi-carrier schemes in BB-PLC. For baseband OFDM systems, signals are transmitted without up-down conversion onto carrier frequency. The complex conjugate property, known as the Hermitian Symmetry Property (HSP) [18], must be satisfied in frequency domain after modulation [19].

Using Fast Fourier Transform (FFT), OFDM significantly enhances the system performance in a dispersive propagation environment [20]. Nevertheless, it is vulnerable to synchronization imperfections. In this work, two related issues caused by a non-ideal sampling clock at the receiver are considered, including:

Chen Chen, Yun Chen, Na Ding, Yizhi Wang, and Xiaoyang Zeng are with the ASIC and System State Key Laboratory, Fudan University, Shanghai, China (e-mail: {10212020002, chenyun, xyzeng}@fudan.edu.cn).

Jia-Chin Lin is with the Department of Communication Engineering, National Central University, Taiwan (e-mail: jiachin@ce.ncu.edu.tw).

Defeng (David) Huang is with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Australia (e-mail: david.huang@uwa.edu.au).

- 1) *Sampling phase offset (SPO)*: It is caused by the misaligned initial sampling instants between the transmitter and the receiver, resulting in an extra phase rotation proportional to the tone index within one OFDM block (*inter-block phase rotations*) [18]. It is equivalent to a fractional symbol timing offset (STO) [21]. Although it could be incorporated into CTF and eliminated by a frequency equalizer (FEQ) [18], the residual part should be compensated due to oscillator instabilities [22].
- 2) *Sampling clock offset (SCO)*: It is caused by a non-synchronized sampling clock between the transmitter and the receiver, triggering not only a inter-block phase rotation, but also a phase growing linearly for successive OFDM blocks (*intra-block phase rotations*) [18].

Both SPO and SCO could be tracked recursively using pilots by a delay-locked-loop (DLL) [22], [23]. The loop operation demands an accurate initial timing acquisition. To the authors' best knowledge, joint sampling timing estimators are seldom discussed in literature. Thus, for comparison, pilot-assisted SPO and SCO estimators are investigated:

- *SPO estimators*: [24] proposed a minimum mean square error (MMSE) estimator for narrowband signals. A low complexity estimator was proposed in [25] taking advantage of the phase rotation between two adjacent pilots. [26] presented several improved variants to better utilize the phase and power characteristics within one or several OFDM blocks. The scheme in [27] further reduced the pilot distortion using the amplitude information and thus enhanced the performance in a multipath channel. Channel state information (CSI) is explicitly required in [24], while implicitly in [25]–[27].
- *SCO estimators*: [28] proposed an estimator using the phase rotation between the upper and the lower frequency band without CSI knowledge. In [29], Morelli et al. proposed a maximum likelihood decoupled estimator (MLDE) and a reduced-complexity estimator (RCE) with superior performance in a multipath channel. CSI could be estimated by the maximum likelihood criterion in [29]. The scheme in [30] was initially devised for MIMO-OFDM, which was modified to suit a single-antenna system in [29]. [31] proposed a weighted joint least square estimator (W-JLSE) outperforming conventional LSE estimators [20]. Both [30] and [31] require CSI.

In the present work, a novel joint acquisition technique for SPO and SCO is proposed. An approximated form of the log-likelihood function is derived under pairwise correlation and joint Gaussian assumption of received signals in frequency domain. To avoid the high complexity two-dimension (2-D) grid-search which cannot be decomposed into a one-dimension (1-D) one due to mutual dependency, a five-step method is proposed. First of all, ancillary phases are obtained block-wise, either by a non-data-aided (NDA) estimator or by its data-aided (DA) counterpart. Then, phase unwrapping restores the actual phases from the wrapped ancillary ones. The resultant phases are further embedded into a linear model. With such linearity, LSE is used to separate SPO and SCO. Bias erasure removes the inherent bias of estimation, followed

by another LSE which leads to refined estimations. The proposed technique highlights the intra-block rotations with performance improvements over conventional pilot-assisted schemes. Effects of various types of noise and violation of the joint Gaussian assumption are analyzed.

The remainder of this paper is structured as follows: Section II introduces the signal model. Section III illustrates the approximated log-likelihood function, several variants of the proposed technique, and analytical expressions for variances and biases. Section IV discusses the applicability to time domain non-Gaussianity, cyclo-stationarity, and temporal correlation. Section V presents simulation results. Finally, Section VI concludes the paper.

## II. SIGNAL MODEL

<sup>1</sup> For  $Q$  OFDM blocks in frequency domain with  $N$  subcarriers in each block, the output signal  $s(t)$  at the transmitter on baseband is [23]

$$s(t) = \sum_{q=0}^{Q-1} \sum_{k \in \mathcal{K}} X_{q,k} \Psi_{q,k}(t) \quad (1)$$

where

$$\Psi_{q,k}(t) = \frac{1}{\sqrt{N}} e^{j2\pi k \left( \frac{t}{NT_{sam}} - \frac{N_g + qN_s}{N} \right)} u(t - qN_s T_{sam}) \quad (2)$$

is the subcarrier pulse,  $\mathcal{K}$  the locations of the data subcarriers,  $X_{q,k}$  the equi-probable data modulated on the  $k$ -th subcarrier in the  $q$ -th OFDM block and  $X_{q,k} \in \mathcal{T}$ ;  $\mathcal{T}$  denotes the finite alphabet of constellation.  $N_s = N + N_g$  is the length of an OFDM block,  $N_g$  the length of the cyclic prefix (CP), and  $T_{sam}$  the sampling interval at the transmitter. The rectangular function  $u(t - qN_s T_{sam})$  is

$$u(t) = \begin{cases} 1 & 0 \leq t \leq N_s T_{sam}, \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The multipath channel is

$$h(t, \tau) = \sum_{l=0}^{L-1} h_l(t) \delta(\tau - \tau_l) \quad (4)$$

where  $L$  denotes the total number of taps,  $\{h_l(t)\}_{l=0,1,\dots,L-1}$  the real channel gains,  $\{\tau_l\}_{l=0,1,\dots,L-1}$  the time delay of each path, and  $\delta(\cdot)$  the delta function. Transmitting  $s(t)$  through the multipath channel (4) yields

$$r(t) = s(t) * h(t, \tau) + w(t) = \sum_{l=0}^{L-1} h_l(t) s(t - \tau_l) + w(t) \quad (5)$$

<sup>1</sup>The following notations are used:  $(\cdot)^*$  as the complex conjugation;  $(\cdot)^T$  the transposition;  $(\cdot)^H$  the Hermitian transposition;  $E\{\cdot\}$  the expectation;  $\text{Var}\{\cdot\}$  the variance;  $\Re\{\cdot\}$  the real part of its argument;  $\Im\{\cdot\}$  the imaginary part of its argument;  $\mathbb{Z}$  the set of all integers;  $\mathbf{I}_N$  the  $N \times N$  the identity matrix;  $\mathbf{J}_N$  the  $N \times N$  anti-diagonal matrix;  $\mathbf{0}_{N_1 \times N_2}$  the  $N_1 \times N_2$  all zero matrix;  $\mathcal{CN}(\mu, \sigma)$  the complex Gaussian distribution with mean  $\mu$  (or mean vector  $\boldsymbol{\mu}$ ) and variance  $\sigma$  (or covariance matrix  $\boldsymbol{\sigma}$ );  $\text{Tr}\{\cdot\}$  the trace of a square matrix;  $\max[\cdot]$  the maximum value of the arguments;  $\text{diag}\{\mathbf{X}\}$  the diagonal matrix with entries given by vector  $\mathbf{X}$ ;  $\boldsymbol{\Sigma}_{\mathbf{X}}$  the autocorrelation matrix of  $\mathbf{X}$  defined as  $E\{\mathbf{X}\mathbf{X}^H\}$ ;  $\mathbf{\Sigma}_{\mathbf{X}}^p$  the pseudocorrelation matrix of  $\mathbf{X}$  defined as  $E\{\mathbf{X}\mathbf{X}^T\}$ ;  $[\cdot]_{i,j}$  the  $(i,j)$ -th entry of a matrix; Other notations would be either self-evident or explained therein.

where  $*$  stands for linear convolution, and  $w(t)$  is the noise with an unknown distribution and correlation.

At the receiver, sampling  $r(t)$  at time instants  $nT'_{sam}$  gives

$$r(nT'_{sam}) = \sum_{l=0}^{L-1} h_l(nT'_{sam})s(nT'_{sam} - \tau_l) + w(nT'_{sam}) \quad (6)$$

where  $T'_{sam}$  is the sampling interval at the receiver. After removing CP, the  $N$  samples in the  $q$ -th OFDM block is denoted as  $\mathbf{r}_q = [r_{q,0}, r_{q,1}, \dots, r_{q,n}, \dots, r_{q,N-1}]$  where  $r_{q,n} = r((n + qN_s + N_g)T'_{sam})$ . However, a timing offset  $\xi$  exists due to timing imperfections, which consists of the integer part  $\text{Int}\{\xi\}$  (caused by the misalignment between the correct starting position of FFT window and estimated position) and the fractional part  $\text{Frac}\{\xi\}$  (caused by the SPO). Without loss of generality, a *left-shift* of FFT window is assumed with  $\text{Int}\{\xi\} \in [-(L-1), 0]$ . Thus,  $r_{q,n} = r((n + qN_s + N_g + \xi)T'_{sam})$ .

The channel is assumed to be quasi-static within one OFDM block's duration. Therefore,  $h_l(q)$  denotes the real channel gain of the  $l$ -th path in the  $q$ -th OFDM block. After demodulation of the  $q$ -th received OFDM block  $\mathbf{r}_q$  by FFT, the  $k$ -th subcarrier is given by [20]

$$R_{q,k} = X_{q,k} H_{q,k} e^{j\Theta_{q,k,\xi,\eta}} \Pi(\eta k) + \sum_{\substack{k'=0 \\ k' \neq k}}^{N-1} \overbrace{X_{q,k'} H_{q,k'} e^{j\Phi(q,k,k',\xi,\eta)} \Pi(k'(1+\eta) - k)}^{\text{ICI}_{\xi,\eta;q,k}} + W_{q,k} \quad (7)$$

where

$$W_{q,k} = \sum_{n=0}^{N-1} w((n + qN_s + N_g + \xi)T'_{sam}) e^{-j2\pi \frac{kn}{N}} \quad (8)$$

$$\Pi(\eta k) = \frac{\sin(\pi \eta k)}{N \sin\left(\frac{\pi \eta k}{N}\right)} \quad (9)$$

$$\Phi(q, k, k', \xi, \eta) = \frac{\pi(N-1)(k'(1+\eta) - k) + 2\pi(1+\eta)\xi k'}{N} + \frac{2\pi q N_s \eta k' + 2\pi N_g \eta k'}{N} \quad (10)$$

$$\Theta_{q,k,\xi,\eta} = \Phi(q, k, k, \xi, \eta) \quad (11)$$

Here,  $H_{q,k} = \sum_{l=0}^{L-1} h_l(q) e^{-j2\pi \frac{k\tau_l}{NT_{sam}}}$  is the CTF of channel [32], and  $\eta = (T'_{sam} - T_{sam})/T_{sam}$  the SCO. The diagram illustrating SPO and SCO is drawn in Fig. 1. The extra phase rotation  $\Theta_{q,k,\xi,\eta}$  can be approximated by

$$\Theta_{q,k,\xi,\eta} \approx \frac{2\pi \xi k}{N} + \pi \eta k + \frac{2\pi(qN_s + N_g)\eta k}{N} \quad (12)$$

for large  $N$  and small  $\eta$ .  $\Pi(\eta k)$  is the amplitude attenuation which approaches 1 for small  $\eta$  [32].  $\text{ICI}_{\xi,\eta;q,k}$  is the inter-carrier interference (ICI) caused by  $\xi$  and  $\eta$ .

For convenience,  $\text{ICI}_{\xi,\eta;q,k}$  is neglected; its effect is demonstrated in Section V. Thus, (7) can be written into a compact matrix form below:

$$\mathbf{R}_q = \Phi_q \mathbf{H}_q \mathbf{X}_q + \mathbf{W}_q \quad (13)$$

where  $\mathbf{R}_q \triangleq [R_{q,0} \ R_{q,1} \ \dots \ R_{q,N-1}]^T$ ,  $\Phi_q \triangleq \text{diag}\{e^{j\Theta_{q,k,\xi,\eta}}, k \in [0, N-1]\}$ ,  $\mathbf{H}_q \triangleq \text{diag}\{\mathbf{H}_q\}$ ,  $\mathbf{H}_q \triangleq [H_{q,0} \ H_{q,1} \ \dots \ H_{q,N-1}]^T$ ,  $\mathbf{X}_q \triangleq [X_{q,0} \ X_{q,1} \ \dots \ X_{q,N-1}]^T$ ,  $\mathbf{W}_q \triangleq [W_{q,0} \ W_{q,1} \ \dots \ W_{q,N-1}]^T$ . Here, we make two assumptions:

$$(A1) \ \mathbf{H}_q \mathbf{X}_q \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_X^2 \sigma_H^2 \mathbf{I}_N).$$

$$(A2) \ \mathbf{W}_q \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_W^2 \mathbf{I}_N) \text{ and independent from } \mathbf{H}_q \mathbf{X}_q.$$

Therefore,  $\mathbf{R}_q$  is a *joint Gaussian vector* distributed as  $\mathcal{CN}(\mathbf{0}_{N \times 1}, (\sigma_X^2 \sigma_H^2 + \sigma_W^2) \mathbf{I}_N)$ . On the other hand, the *pseudocorrelation matrix* (PCM) of  $\mathbf{H}_q \mathbf{X}_q$  and  $\mathbf{W}_q$  can be written into

$$[\Sigma_{\mathbf{H}\mathbf{X}}^p]_{i,j} = \begin{cases} \sigma_X^2 \sigma_H^2, & i+j = N \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

$$[\Sigma_{\mathbf{W}}^p]_{i,j} = \begin{cases} \sigma_W^2, & i+j = N \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

using the proprieties and HSPs of  $\mathbf{H}_q \mathbf{X}_q$  and  $\mathbf{W}_q$  respectively. For instance, the propriety of  $W_{q,k}$  indicates that  $E\{W_{q,k} W_{q,k'}\} = 0$  if  $k \neq N-k'$ , while the HSP implies that  $W_{q,k} = W_{q,N-k}^*$  and so  $E\{W_{q,k} W_{q,k'}\} = \sigma_W^2$  if and only if  $k = N-k'$ .

*Remarks:*

In practice, neither  $X_{q,k}$  nor  $H_{q,k}$  is Gaussian distributed:  $X_{q,k}$  is taken from a finite alphabet with equal probability and thus is uniformly distributed, while the statistical model in [6] is adopted for  $H_{q,k}$ . The closed-form distribution of  $\mathbf{H}_q \mathbf{X}_q$  is hard to pursue in practice. Even when both  $X_{q,k}$  and  $H_{q,k}$  are Gaussian distributed, their product  $H_{q,k} X_{q,k}$  follows the Gaussian product distribution which is non-Gaussian. (A1) and (A2) significantly simplify our analysis by making  $R_{q,k} \sim \mathcal{CN}(0, \sigma_X^2 \sigma_H^2 + \sigma_W^2)$ . Robustness of the proposed schemes against non-Gaussianity of  $\mathbf{R}_q$  is demonstrated via simulations in Section V.

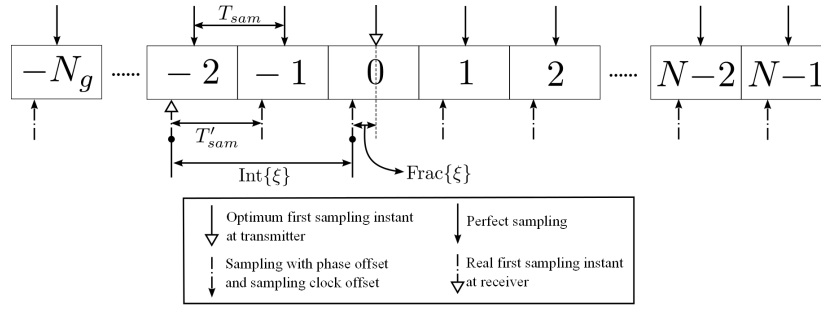
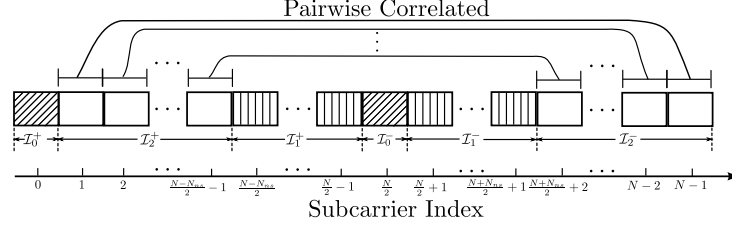
### III. PROPOSED ESTIMATORS

#### A. Approximated Log-likelihood Function

An OFDM system with  $N_n$  null subcarriers at both ends of the spectrum as the guard bands is considered here. The *universal set* containing all the  $N$  subcarriers is denoted as  $\mathcal{I}_U$ , which can be decomposed into three subsets with index given by

- $\mathcal{I}_0 = \{k | k = 0, k = N/2, k \in \mathbb{Z}\} = \mathcal{I}_0^+ \cup \mathcal{I}_0^-$
- $\mathcal{I}_1 = \{k | k = (N - N_n)/2 + v, 0 \leq v \leq (N_n/2) - 1, v \in \mathbb{Z}\} \cup \{k | k = (N/2) + v, 1 \leq v \leq (N_n + 1)/2 + 1, v \in \mathbb{Z}\} = \mathcal{I}_1^+ \cup \mathcal{I}_1^-$
- $\mathcal{I}_2 = \{k | 1 \leq k \leq (N - N_n)/2 - 1, k \in \mathbb{Z}\} \cup \{k | (N - N_n)/2 + 2 \leq k \leq N - 1, k \in \mathbb{Z}\} = \mathcal{I}_2^+ \cup \mathcal{I}_2^-$

For BB-PLC baseband system,  $\mathcal{I}_0$  delivers random *real-valued* data, which are often nullified since direct current (DC) and Nyquist frequency subcarriers are generally discarded.  $\mathcal{I}_1$  contains the null subcarriers, while subcarriers in  $\mathcal{I}_2$  transmit data. Each subset can be divided into the left and the right half. The cardinalities for  $\mathcal{I}_0 \sim \mathcal{I}_2$  are:  $\text{Card}\{\mathcal{I}_0\} = 2$ ,  $\text{Card}\{\mathcal{I}_1\} = N_n$ ,  $\text{Card}\{\mathcal{I}_2\} = N_u \triangleq N - N_n - 2$

Fig. 1. Diagram for SPO, SCO,  $\text{Int}\{\xi\}$ , and  $\text{Frac}\{\xi\}$  for the  $q$ -th OFDM block.Fig. 2. Decomposition into  $\mathcal{I}_0 = \mathcal{I}_0^+ \cup \mathcal{I}_0^-$ ,  $\mathcal{I}_1 = \mathcal{I}_1^+ \cup \mathcal{I}_1^-$ , and  $\mathcal{I}_2 = \mathcal{I}_2^+ \cup \mathcal{I}_2^-$ .

For the  $q$ -th block, an alternative sequence can be formulated by taking the conjugation of the right half of  $\mathbf{R}_q$ :

$$\mathbf{R}'_q \triangleq \left[ \underbrace{R_{q,0} \cdots R_{q,N/2-1}}_{N/2} \underbrace{R_{q,N/2}^* \cdots R_{q,N-1}^*}_{N/2} \right]^T \quad (16)$$

Now, the *autocorrelation matrix* of  $\mathbf{R}'_q$  could be written into

$$\Sigma_{\mathbf{R}'_q} = \begin{bmatrix} \sigma_W^2 & \mathbf{0}_{1 \times (N-1)} \\ \mathbf{0}_{(N-1) \times 1} & \mathbf{B} \end{bmatrix} \quad (17)$$

where the sub-matrix  $\mathbf{B}$  is shown in (18) on the next page.  $\mathbf{C}_{N_n+1}$  represents the  $(N_n+1) \times (N_n+1)$  bi-diagonal matrix with unity diagonal and anti-diagonal entries<sup>2</sup>, and  $\Theta_{q,\theta} = \Theta_{q,k,\xi,\eta} + \Theta_{q,N-k,\xi,\eta} = 2\pi\xi + \eta(\pi N + 2\pi N_g + 2\pi q N_s) \forall k$  where  $\theta = [\xi \ \eta]^T$ .  $\mathbf{B}_n$  is closely tied with the noise structure since  $\mathbf{B}_n = [\Sigma_{\mathbf{W}}]_{1:N-1,1:N-1} + [\Sigma_{\mathbf{W}}^p]_{1:N-1,1:N-1}$  where  $[\cdot]_{N_1:N_2,N_3:N_4}$  represents the sub-matrix spanning rows  $N_1 \sim N_2$  and columns  $N_3 \sim N_4$ . Discussions of  $\mathbf{B}_n$  under various scenarios are given in Section IV.

Arranging observations of  $Q$  OFDM blocks into a vector  $\mathbf{R}' \triangleq [\mathbf{R}'_0 \ \mathbf{R}'_1 \ \mathbf{R}'_2 \ \cdots \ \mathbf{R}'_{Q-1}]$ , after manipulations reported in Appendix A, the *approximated log-likelihood function* conditioned on  $\theta$  under (A1) and (A2) is

$$\Lambda(\mathbf{R}'|\theta) = \text{Const.} + 2 \sum_{q=0}^{Q-1} \sum_{k \in \mathcal{I}_2^+} \Re\{\lambda_1(q,k) e^{-j\Theta_{q,\theta}}\} \quad (19)$$

$$\lambda_1(q,k) = R_{q,k} R_{q,N-k} \quad (20)$$

Const. is irrelevant to estimation and can be neglected.  $\Lambda(\mathbf{R}'|\theta)$  coherently accumulates  $\Re\{\lambda_1(q,k) e^{-j\Theta_{q,\theta}}\}$  across all  $Q$  OFDM blocks and subcarriers in  $\mathcal{I}_2$ .

<sup>2</sup>For instance,  $\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ,  $\mathbf{J}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ ,  $\mathbf{C}_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

### B. Joint Estimation of $\xi$ and $\eta$

Joint estimation of  $\theta$  using (19) takes the form

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta} \Lambda(\mathbf{R}'|\theta) \\ &\triangleq \arg \max_{\theta} \left( \sum_{q=0}^{Q-1} \sum_{k \in \mathcal{I}_2^+} \Re\{\lambda_1(q,k) e^{-j\Theta_{q,\theta}}\} \right) \end{aligned} \quad (21)$$

where  $\hat{\theta}$  is the estimation of  $\theta$ . (21) requires a 2-D grid search on all possible  $\theta$  for the global maxima. The complexity involved can be formidably high and cannot be pursued in practice [29]. One seemingly plausible solution is to decouple the original 2-D problem into a 1-D one as in [29]. However, it is not possible since estimations of  $\xi$  and  $\eta$  are *mutually dependent*: estimating  $\xi$  by fixing  $\eta$  introduces bias into  $\hat{\xi}$  and vice versa. A five-step method is employed to overcome this issue:

- 1) **Ancillary Phases Estimation:** Obtaining ancillary phases  $\hat{\Theta}_{q,\theta}, \forall q \in [0, Q-1]$  by either a NDA or a DA estimator. For the  $q$ -th OFDM block,

$$\hat{\Theta}_{q,\theta} = \begin{cases} \arg \left\{ \sum_{k \in \mathcal{I}_2^+} \lambda_1(q,k) \right\}, & \text{NDA} \\ \arg \left\{ \sum_{k \in \mathcal{I}_2^+} \lambda_1(q,k) \lambda_2(q,k) \right\}, & \text{DA} \end{cases} \quad (22)$$

where  $\lambda_2(q,k)$  is introduced to highlight the modulation effect and frequency selectivity as

$$\lambda_2(q,k) = [2 + \text{SNR}^{-1}(q,k)]^{-1} \quad (23)$$

where  $\text{SNR}(q,k) = \frac{|X_{q,k}|^2 |H_{q,k}|^2}{\sigma_W^2}$ . This is explained in more details in Appendix A. For notational convenience,  $\mathcal{A}$  denotes either NDA or DA. The DA suits the closed-loop system with feedback to compute  $\lambda_2(q,k)$ . On the

$$\mathbf{B} = \underbrace{\begin{bmatrix} \sigma_X^2 \sigma_H^2 \mathbf{I}_{N_u/2} & \mathbf{0}_{N_u/2 \times (N_n+1)} & (e^{j\Theta_{q,\theta}} \sigma_X^2 \sigma_H^2) \mathbf{J}_{N_u/2} \\ \mathbf{0}_{(N_n+1) \times N_u/2} & \mathbf{0}_{N_n+1} & \mathbf{0}_{(N_n+1) \times N_u/2} \\ (e^{j\Theta_{q,\theta}} \sigma_X^2 \sigma_H^2) \mathbf{J}_{N_u/2} & \mathbf{0}_{N_u/2 \times (N_n+1)} & \sigma_X^2 \sigma_H^2 \mathbf{I}_{N_u/2} \end{bmatrix}}_{\mathbf{B}_s} + \underbrace{\begin{bmatrix} \sigma_W^2 \mathbf{I}_{N_u/2} & \mathbf{0}_{N_u/2 \times (N_n+1)} & \sigma_W^2 \mathbf{J}_{N_u/2} \\ \mathbf{0}_{(N_n+1) \times N_u/2} & \sigma_W^2 \mathbf{C}_{N_n+1} & \mathbf{0}_{(N_n+1) \times N_u/2} \\ \sigma_W^2 \mathbf{J}_{N_u/2} & \mathbf{0}_{N_u/2 \times (N_n+1)} & \sigma_W^2 \mathbf{I}_{N_u/2} \end{bmatrix}}_{\mathbf{B}_n} \quad (18)$$

other hand, the NDA, with a similar form to that in [33], reduces the complexity without calculating  $\lambda_2(q, k)$ .

- 2) **Phase Unwrapping:** Since  $\Theta_{q+1,\theta} - \Theta_{q,\theta} = \text{Const.}$ ,  $\{\Theta_{q,\theta}, q = 0, 1, \dots, Q-1\}$  lie on the same line. However,  $\arg$  in (22) folds  $\widehat{\Theta}_{q,\theta}^A$  into  $[-\pi, +\pi]$  and breaks the line into several discrete segments, known as the phase ambiguity. To deal with this issue, phase unwrapping is used to detect the discontinuities in phases sequentially and restore the actual phases. More specifically, once the criterion

$$|\widehat{\Theta}_{q+1,\theta}^A - \widehat{\Theta}_{q,\theta}^A| \geq \varphi\pi \quad (24)$$

is satisfied,  $\pm 2p\pi$  is added to the phases starting from  $\widehat{\Theta}_{q+1,\theta}^A$ , until (24) is met again (and  $\pm 2(p+1)\pi$  is added afterwards).  $\varphi \in [0, 2]$  is the tolerance level of phase unwrapping.  $p$  denotes the number of times that (24) is met before  $\widehat{\Theta}_{q,\theta}^A$ . The sign in (24) is determined by the direction of phase alternation.

- 3) **Linearization and Least Square Estimation:** The unwrapped phase is denoted as  $\widehat{\Theta}_{q,\theta}^A$ . They could be combined into the vector  $\widetilde{\Theta}_A$ . Now, assuming correct phase unwrapping across all the  $Q$  OFDM blocks,  $\widetilde{\Theta}_A$  can be embedded into a linear model shown as

$$\widetilde{\Theta}_A = \mathbf{E}\theta_A + \mathbf{V} \quad (25)$$

The  $Q \times 2$  observation matrix  $\mathbf{E}$  can be expressed by

$$\mathbf{E} = \begin{bmatrix} 2\pi & 2\pi & \cdots & 2\pi & \cdots & 2\pi \\ D_0 & D_1 & \cdots & D_q & \cdots & D_{Q-1} \end{bmatrix}^T \quad (26)$$

where  $D_q = \pi N + 2\pi N_g + 2\pi q N_s$  and  $\mathbf{V}$  is the self-noise with mean vector  $\mu_{\mathbf{V}}^{\theta, \text{SNR}}$  and covariance matrix  $\Sigma_{\mathbf{V}}^{\theta, \text{SNR}}$ , expressed by

$$\mu_{\mathbf{V}}^{\theta, \text{SNR}} = \left\{ -\frac{\sin(\Theta_{q,\theta})}{\cos(\Theta_{q,\theta}) + \text{SNR}}, 0 \leq q \leq Q-1 \right\} \quad (27)$$

$$\Sigma_{\mathbf{V}}^{\theta, \text{SNR}} = \text{diag} \left\{ \frac{1 - \cos(\Theta_{q,\theta})}{\Xi N \times \text{SNR}}, 0 \leq q \leq Q-1 \right\} \quad (28)$$

which are derived under (A2) and (A3), (A4). (A3), (A4) are shown below:

- (A3) flat fading ( $L = 1$ ) with real channel gain  $h_0(q)$  with zero mean and variance  $\sigma_H^2$ . Thus, in frequency domain,  $\mathbf{E}\{H_{q,k}\} = 0$  and  $\mathbf{E}\{|H_{q,k}|^2\} = \sigma_H^2$ .
- (A4) Constant modulus modulated signal satisfying  $\mathbf{E}\{X_{q,k}\} = 0$  and  $\mathbf{E}\{|X_{q,k}|^2\} = \sigma_X^2$ .

Notice that, (A3) and (A4) differ from (A1): they merely impose the finite first moment and second moment on  $X_{q,k}$  and  $H_{q,k}$  without the Gaussian assumption on  $\mathbf{H}_q \mathbf{X}_q$ .  $\Xi = \frac{N_s}{2N}$  is the ratio between usable pairwise correlations to the useful FFT duration, and  $\text{SNR} = \nu \frac{\sigma_X^2 \sigma_H^2}{\sigma_W^2}$  where  $\nu = \frac{N_s}{N}$  compensates the loss of discarding unused subcarriers. In lack of  $\theta$ , it is impossible to pre-compute  $\mu_{\mathbf{V}}^{\theta, \text{SNR}}$  and  $\Sigma_{\mathbf{V}}^{\theta, \text{SNR}}$ . An ordinary least square (OLS) estimation yields the initial  $\widehat{\theta}_A$ :

$$\widehat{\theta}_A = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \widetilde{\Theta}_A = \theta_A + \mathbf{V} \quad (29)$$

where  $\mathbf{V} = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \mathbf{V}$  is the  $2 \times 1$  error vector associated with  $\mathbf{V}$  by a linear transformation  $(\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}$ .

- 4) **Bias Erasure:** With  $\widehat{\theta}_A$  and assuming perfect a priori knowledge of SNR, both  $\mu_{\mathbf{V}}^{\theta, \text{SNR}}$  and  $\Sigma_{\mathbf{V}}^{\theta, \text{SNR}}$  can be replaced by  $\mu_{\mathbf{V}}^{\widehat{\theta}_A, \text{SNR}}$  and  $\Sigma_{\mathbf{V}}^{\widehat{\theta}_A, \text{SNR}}$ . Bias erasure (BE) of  $\mu_{\mathbf{V}}^{\widehat{\theta}_A, \text{SNR}}$  in (25) gives

$$\underline{\theta}_A = \widetilde{\Theta}_A - \mu_{\mathbf{V}}^{\widehat{\theta}_A, \text{SNR}} = \mathbf{E}\theta_A + \mathbf{V}' \quad (30)$$

where  $\mu_{\mathbf{V}'}^{\widehat{\theta}_A, \text{SNR}} = \mu_{\mathbf{V}}^{\theta_A, \text{SNR}} - \mu_{\mathbf{V}}^{\widehat{\theta}_A, \text{SNR}}$  and  $\Sigma_{\mathbf{V}'}^{\widehat{\theta}_A, \text{SNR}} = \Sigma_{\mathbf{V}}^{\widehat{\theta}_A, \text{SNR}}$ .

- 5) **Re-estimation:** Updating the estimation using the weighted least square estimation after bias erasure (WLS-BE) by

$$\underline{\theta}_A|_{\text{WLS-BE}} = (\mathbf{E}^T [\Sigma_{\mathbf{V}'}^{\widehat{\theta}_A, \text{SNR}}]^{-1} \mathbf{E})^{-1} \mathbf{E}^T [\Sigma_{\mathbf{V}'}^{\widehat{\theta}_A, \text{SNR}}]^{-1} \underline{\theta}_A \quad (31)$$

or OLS after bias erasure (OLS-BE) ignoring the weight factors

$$\underline{\theta}_A|_{\text{OLS-BE}} = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \underline{\theta}_A \quad (32)$$

The performance loss of (32) comparing with (31) could be neglected, since  $\Sigma_{\mathbf{V}'}^{\widehat{\theta}_A, \text{SNR}}$  approaches a diagonal matrix with nearly identical entries as  $\cos(\Theta_{q,\theta})$  varies slowly block-wise given small  $\eta$ .

### C. Variance and Bias of Estimation

$\text{Var}\{\widehat{\xi}\}$ ,  $\text{Var}\{\widehat{\eta}\}$ ,  $\text{Bias}\{\widehat{\xi}\}$ , and  $\text{Bias}\{\widehat{\eta}\}$  before BE with perfect phase unwrapping can be computed straightforwardly from (25) under (A2), (A3), and (A4). The details are shown in Appendix B. Analytical expressions of  $\text{Var}\{\widehat{\xi}\}$ ,  $\text{Var}\{\widehat{\eta}\}$ ,  $\text{Bias}\{\widehat{\xi}\}$ , and  $\text{Bias}\{\widehat{\eta}\}$  are given by (33), (34), (35), (36) respectively.

$$\text{Var}\{\hat{\xi}\} = \frac{2g^2(2Q+1)(Q+1) + 2g(Q+1)(4Q-1) + (2Q+1)(2Q-1)}{4\pi^2 N(1+g)^2 Q(Q^2-1)\Xi\text{SNR}} - \frac{\sum_{q=0}^{Q-1} \cos(\Theta_{q,\theta}) \left[ (Q-1)[1+4g+4Q(1+g)] - 6q[g+Q(1+g)] \right]^2}{4\pi^2(1+g)^2 N Q^2(Q+1)^2(Q-1)^2 \Xi\text{SNR}} \quad (33)$$

$$\text{Var}\{\hat{\eta}\} = \frac{3}{\pi^2 N^3(1+g)^2 Q(Q^2-1)\Xi\text{SNR}} - \frac{\sum_{q=0}^{Q-1} 9 \cos(\Theta_{q,\theta})(2q-Q+1)^2}{\pi^2(1+g)^2 N^3 Q^2(Q+1)^2(Q-1)^2 \Xi\text{SNR}} \quad (34)$$

$$\text{Bias}\{\hat{\xi}\} = - \sum_{q=0}^{Q-1} \frac{\sin(\Theta_{q,\theta})}{\cos(\Theta_{q,\theta}) + \text{SNR}} \left[ \frac{(Q-1)[3+6g+2(1+g)(2Q-1)] - 6q[1+2g+(1+g)(Q-1)]}{2\pi(1+g)Q(Q+1)(Q-1)} \right] \quad (35)$$

$$\text{Bias}\{\hat{\eta}\} = \sum_{q=0}^{Q-1} \frac{\sin(\Theta_{q,\theta})}{\cos(\Theta_{q,\theta}) + \text{SNR}} \left[ \frac{3(Q-1) - 6q}{\pi(1+g)NQ(Q+1)(Q-1)} \right] \quad (36)$$

Here,  $g = \frac{N_g}{N}$  is the ratio between CP and useful FFT duration with details in Appendix B. The mean square errors (MSEs) are

$$\text{MSE}\{\hat{\xi}\} = \text{Var}\{\hat{\xi}\} + [\text{Bias}\{\hat{\xi}\}]^2 \quad (37)$$

$$\text{MSE}\{\hat{\eta}\} = \text{Var}\{\hat{\eta}\} + [\text{Bias}\{\hat{\eta}\}]^2 \quad (38)$$

With perfect BE,  $\mu_{\mathbf{V}}^{\theta_{\mathbf{A}}, \text{SNR}} = \mathbf{0}_{Q \times 1}$ , and consequently the biases in (35) and (36) vanish which makes  $\text{MSE}\{\hat{\xi}\} = \text{Var}\{\hat{\xi}\}$  and  $\text{MSE}\{\hat{\eta}\} = \text{Var}\{\hat{\eta}\}$ .

*Remarks:*

(i): The first terms of (33) and (34) are written as  $\text{Var}\{\hat{\xi}\}_1$  and  $\text{Var}\{\hat{\eta}\}_1$ . It is easy to prove that  $\max[\text{Var}\{\hat{\xi}\}] = 2\text{Var}\{\hat{\xi}\}_1$  and  $\max[\text{Var}\{\hat{\eta}\}] = 2\text{Var}\{\hat{\eta}\}_1$ . They provide a good criterion on the choice of  $Q$  and  $N$ . For instance, when  $N$  is fixed, to warrant the performance of  $\text{MSE}\{\hat{\xi}\}_{\text{Target}}$  and  $\text{MSE}\{\hat{\eta}\}_{\text{Target}}$  under a certain SNR assuming perfect BE,  $Q$  should be chosen according to (39) shown on the next page, for  $Q \gg 1$  by setting  $\text{MSE}\{\hat{\xi}\}_{\text{Target}} = \max[\text{Var}\{\hat{\xi}\}]$  and  $\text{MSE}\{\hat{\eta}\}_{\text{Target}} = \max[\text{Var}\{\hat{\eta}\}]$ ;  $\Omega$  denotes the SNR gap considering the possible loss of accuracy of (33) and (34).

(ii): Obviously,  $\lambda_2(q, k)$  in (23) reduces to  $\text{SNR}(q, k)$  in low SNR, and becomes 0.5 in high SNR. Thus, the advantage of DA is more significant in low SNR.

(iii): For correct phase unwrapping in absence of noise, two conditions must be satisfied:

$$(C1) \quad -\pi \leq 2\pi\xi + \pi\eta N + 2\pi N_g \eta \leq +\pi \text{ for correct } \xi \text{ estimation}$$

$$(C2) \quad -2\pi \leq 2\pi N_g \eta \leq 2\pi \text{ for correct } \eta \text{ estimation}$$

Note that, only  $\text{Frac}\{\xi\}$  is estimable, since  $\text{Int}\{\xi\}$  results in an imperceptible  $2\pi\text{Int}\{\xi\}$  rotation in  $\Theta_{q,\theta}$ ; the only impact of  $\text{Int}\{\xi\}$  on  $\text{Frac}\{\xi\}$  is an increased  $\text{ICI}_{\xi,\eta;q,k}$ . Thus,  $\xi$  is treated equivalently as  $\text{Frac}\{\xi\}$  and bounded in  $[-0.5, 0.5]$ . The *hexagon* denoting the solution area is plotted in Fig. 3, with area  $S = \frac{3}{2(N+N_g)}$ , which shrinks in presence of noise.

(iv): In practice,  $\lambda_2(q, k)$  (for the DA scheme) and SNR (for both DA and NDA schemes) need to be calculated. Firstly,

$\lambda_2(q, k)$  in (23) is computed as

$$\widehat{\lambda_2(q, k)} = \left[ 2 + \overline{\text{SNR}^{-1}(q, k)} \right]^{-1} \quad (40)$$

where  $\overline{\text{SNR}^{-1}(q, k)} = \left( \frac{|\overline{X_{q,k}}|^2 |\widehat{H_{q,k}}|^2}{\sigma_W^2} \right)^{-1}$ .  $\overline{X_{q,k}}$  is either  $X_{q,k}$  (for pilots or null subcarriers), or the output of decision device  $\widehat{X_{q,k}}$ .  $\widehat{H_{q,k}}$  stands for the estimated CTF. With (18),  $\sigma_W^2$  and SNR can be estimated by

$$\widehat{\sigma_W^2} = \frac{2 \sum_{q=0}^{Q-1} \sum_{k \in \mathcal{I}_1^+} \lambda_1(q, k)}{QN_n} \quad (41)$$

$$\widehat{\text{SNR}} = \nu \frac{\widehat{\sigma_X^2} \widehat{\sigma_H^2}}{\widehat{\sigma_W^2}} = \nu \left( \frac{\widehat{\sigma_X^2}}{\widehat{\sigma_W^2}} - 1 \right) \quad (42)$$

where

$$\widehat{\sigma_X^2} \triangleq \widehat{\sigma_X^2} \widehat{\sigma_H^2} + \widehat{\sigma_W^2} = \frac{2 \sum_{q=0}^{Q-1} \sum_{k \in \mathcal{I}_2^+} |R_{q,k}|^2}{QN_u} \quad (43)$$

Now,  $\widehat{\text{SNR}}$  is inserted back into (27) and (28) to yield  $\mu_{\mathbf{V}}^{\theta_{\mathbf{A}}, \widehat{\text{SNR}}}$ . Replacing  $\mu_{\mathbf{V}}^{\theta_{\mathbf{A}}, \text{SNR}}$  in (30) with  $\mu_{\mathbf{V}}^{\theta_{\mathbf{A}}, \widehat{\text{SNR}}}$  leads to  $\underline{\theta}_{\mathbf{A}}' = \underline{\theta}_{\mathbf{A}} - \mu_{\mathbf{V}}^{\theta_{\mathbf{A}}, \widehat{\text{SNR}}}$ . Using OLS-BE in (32) gives the practical estimation  $\underline{\theta}_{\mathbf{A}}|_{\text{PS}}$ , termed as the *practical scheme* (PS) given by:

$$\underline{\theta}_{\mathbf{A}}|_{\text{PS}} = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \underline{\theta}_{\mathbf{A}}' \quad (44)$$

#### IV. APPLICABILITY TO DIFFERENT TYPES OF NOISE

##### A. Class-A Impulsive Noise

The PDF for time domain real Class-A noise random variable (r.v.)  $w$  is shown in Table I [12].  $\mathcal{F}$  denotes the FFT operation.  $T = \sigma_g^2 / \sigma_i^2$  is the power ratio between the Gaussian component with variance  $\sigma_g^2$  and the additional man-made noise component with variance  $\sigma_i^2$ . The Class-A model combines both components and its PDF can be regarded as a weighted sum of infinitely many Gaussian PDFs with an increasing variance  $\sigma_p^2$ . Impulsive index  $A$  is the product of the average rate of impulsive noise and the mean duration of a typical impulse. The noise is *more impulsive* for a small  $A$ , whereas for  $A \rightarrow \infty$ , the PDF becomes

$$Q = \max \left[ \frac{2}{N} \left( \text{MSE} \{ \hat{\xi} \}_{\text{Target}} \pi^2 \Xi (\text{SNR} + \Omega) \right)^{-1}, \frac{\sqrt[3]{6}}{N} \left( \text{MSE} \{ \hat{\eta} \}_{\text{Target}} \pi^2 (1+g)^2 \Xi (\text{SNR} + \Omega) \right)^{-\frac{1}{3}} \right] \quad (39)$$

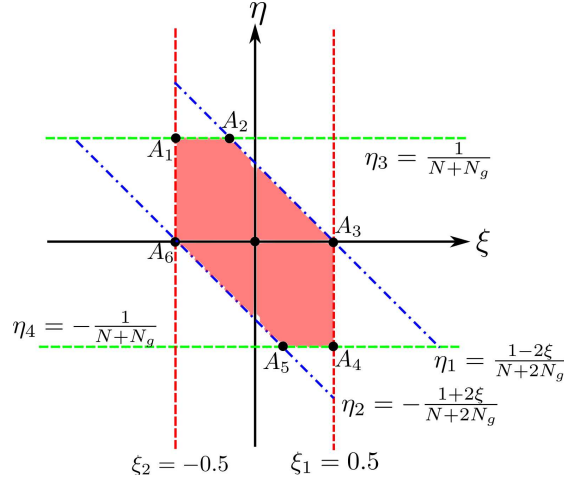


Fig. 3. The region of  $\xi$  and  $\eta$  for perfect phase unwrapping in a noiseless environment. The coordinates for  $A_1 \sim A_6$  are:  $A_1 : (-0.5, (N+N_g)^{-1})$ ,  $A_2 : (-N_g(2(N+N_g))^{-1}, (N+N_g)^{-1})$ ,  $A_3 : (0.5, 0)$ ,  $A_4 : (0.5, (N+N_g)^{-1})$ ,  $A_5 : (N_g(2(N+N_g))^{-1}, -(N+N_g)^{-1})$ ,  $A_6 : (-0.5, 0)$ .

TABLE I  
STATISTICAL INFORMATION FOR CLASS-A AND NAKAGAMI-M DISTRIBUTION

Noise	Parameters	PDF (time domain)	PDF (frequency domain)
Class-A	$A, T$	$f(w) = \sum_{p=0}^{\infty} \frac{\alpha_p}{\sqrt{2\pi}\sigma_p} \exp(-\frac{w^2}{2\sigma_p^2})$	$W = \mathcal{F}\{w\}$
		$\alpha_p = e^{-A} \frac{A^p}{p!}$ $\sigma_p^2 = \sigma_g^2 \frac{p/A+T}{T}$	$W \sim \mathcal{CN}(0, \frac{T+1}{T}\sigma_g^2)$
Nakagami-m	$m, \Omega$	$f(u) = \frac{2m^m(u)^{2m-1}}{\Gamma(m)\Omega^m} e^{-\frac{mu^2}{\Omega}}$	$W = \mathcal{F}\{\Re\{w\}\}$
		$m = \frac{\Omega^2}{E\{(u^2-\Omega)^2\}}$	$W \sim \mathcal{CN}(0, \Omega)$
		$\Omega = E\{u^2\}$ $\Re\{w\} = \sqrt{2}u \cos(\zeta), \zeta \sim \mathcal{U}(-\pi, +\pi)$	

a Gaussian distribution [14]. After FFT,  $W$  in frequency domain approaches Gaussianity, particularly for a large  $A$ , which is accurate for  $p > 3$  [13].

### B. Nakagami-m Background Noise

The envelope PDF of time domain Nakagami-m background noise r.v.  $u$  is given in Table I [10], where  $m$  is the *fading figure* controlling the severity of the amplitude fading [34] and  $m \geq 0.5$ .  $\Gamma(\cdot)$  is the Gamma function, and  $\Omega$  the *second moment* of  $u$  [35]. For the corresponding complex random noise  $w$  with an envelope following the Nakagami-m PDF ( $|w| = u$ ), the *axis-PDF* for its real part  $\Re\{w\} = u \cos(\zeta)$  is shown in power series in [2] and in closed-form in [10];  $\zeta$  is uniformly distributed in  $[-\pi, +\pi]$ . Yet, the closed-form distribution for the corresponding frequency domain r.v.  $W$  is hard to pursue. As a rule of thumb,  $W$  can be modeled as a Gaussian r.v. with zero mean and variance  $\Omega$ , by virtue of the central limit theorem and the uniformly distributed  $\zeta$  [2]. An additional  $\sqrt{2}$  is inserted to compensate the loss of ignoring  $\Im\{w\}$ .

*Remarks:*

(i): In a strict sense,  $m$  differs in different frequency ranges.

For high frequencies,  $m \approx 1$ , while  $m < 1$  for low frequencies [2]. For simplicity, a common  $m$  for all subcarriers is assumed.

(ii): A good tool to evaluate the goodness-of-fit is the  $\chi^2$ -test. In this paper, the D'Agostino's  $\chi^2$ -test is adopted with a significance level  $\alpha = 0.05$ , which is based on the transformation of the third and fourth-order statistics (skewness and kurtosis) on data samples [36].  $T$  is set as  $\{1, 0.1, 0.01\}$  with respect to light, moderate, and heavy impulsiveness, while  $A \in \{0.01, 0.1, 1\}$ . For the Nakagami-m background noise,  $m \in \{0.5, 1.0, 2.0, 10\}$ . The kurtosis and skewness for the Gaussian case are 3 and 0 respectively.

Independent and identically distributed (i.i.d) noise samples are generated using the corresponding PDFs. The averaged in-phase p-value (probability that the distribution cannot be rejected as Gaussian) on usable subcarriers with  $N = 512$  is shown in Table II<sup>3</sup>. Also, the kurtosis and skewness are tabulated. Similar results can be obtained for the quadrature. For most cases, near-Gaussian performances can be achieved. Thus,  $\Sigma_{\mathbf{W}}$ ,  $\Sigma_{\mathbf{W}}^p$ , and  $\mathbf{B}_n$  in (18) are invariant. The Gaussianity

<sup>3</sup>For the Nakagami-m background noise, the DC subcarrier could reach a considerably high value which would destroy the Gaussianity. Fortunately, it is usually discarded in most OFDM systems.

TABLE II  
KURTOSIS, SKEWNESS, AND P-VALUE OF THE D'AGOSTINO'S  $\chi^2$ -TEST FOR CLASS-A AND NAKAGAMI-M NOISE ( $N = 512, \alpha = 0.05$ )

Noise	Parameters		Kurt.(in-phase)	Skew.(in-phase)	p-value(in-phase)
Class-A	$T = 0.01$	$A = 0.01$	<b>2.61</b>	$9.12 \times 10^{-4}$	<b>0.5929</b>
		$A = 0.1$	2.94	$7.21 \times 10^{-4}$	0.9532
		$A = 1$	2.98	$2.10 \times 10^{-3}$	0.9521
	$T = 0.1$	$A = 0.01$	<b>2.69</b>	$-2.7 \times 10^{-3}$	<b>0.7304</b>
		$A = 0.1$	2.94	$-1.3 \times 10^{-4}$	0.9544
		$A = 1$	2.98	$8.75 \times 10^{-4}$	0.9498
	$T = 1$	$A = 0.01$	2.90	$6.78 \times 10^{-4}$	0.9435
		$A = 0.1$	2.98	$1.90 \times 10^{-3}$	0.9502
		$A = 1$	2.98	$-8.1 \times 10^{-4}$	0.9461
Nakagami-m	$m = 0.5$		2.98	$-8.3 \times 10^{-4}$	0.9154
	$m = 1.0$		2.99	$-1.4 \times 10^{-3}$	0.9173
	$m = 2.0$		2.98	$-2.4 \times 10^{-4}$	0.9116
	$m = 10.0$		2.98	$-1.1 \times 10^{-3}$	0.9112

degrades for  $A = 0.01$  with  $T = 0.01$  or  $T = 0.1$ .

### C. Effect of Temporal Correlation and Cyclo-stationarity

For colored noise, the zero vectors in  $\mathbf{B}_n$  are replaced by non-zero values, and both the diagonal and the anti-diagonal elements of  $\mathbf{B}_n$  are not uniform and this degrades the estimation. The cyclo-stationarity could be analyzed in a similar way. The uniformity of the (anti-)diagonal elements in  $\mathbf{B}_n$  is *preserved* which entails no performance loss. Meanwhile, non-zero elements appear in the original zero vectors in  $\mathbf{B}_n$ . As the zero vectors in  $\mathbf{B}_n$  are not involved in the estimation, they would not degrade the performance.

## V. SIMULATION

An OFDM system with the following parameters is considered, unless otherwise mentioned:

- 1) **Sampling interval:**  $T_{sam} = 10ns$  (sampling frequency 100 MHz)
- 2) **Number of subcarriers:**  $N = 512$
- 3) **Guard interval:**  $N_g = 64$
- 4) **OFDM blocks:**  $Q = 10$
- 5) **Null subcarriers:**  $N_n = 64$
- 6) **Modulation:** 16-PSK ( $M = 16, \sigma_X^2 = 1$ )
- 7) **Channel:** The real channel taps are generated from the model proposed in [6]. Channel classes 1 ~ 9 with different channel capacities are selected with respect to the probability profile  $\{0.0349, 0.1678, 0.1818, 0.1188, 0.1188, 0.1258, 0.0979, 0.0769, 0.0769\}$ . The number of paths  $L$  follows a Poisson distribution. Also, the channel length is tailored to satisfy  $L \leq \frac{N_g}{2} = 32$ . Other parameters are chosen according to [6].
- 8) **Impulsive parameters:**  $A, T \in \{0.01, 0.1, 1\} \times \{0.01, 0.1, 1\}$
- 9) **Nakagami-m fading figure:**  $m \in \{0.5, 1.0, 2.0, 10.0\}$
- 10) **SPO:**  $\xi = 0.1$
- 11) **SCO:**  $\eta = 1 \times 10^{-5}$
- 12) **Phase unwrapping tolerance level:**  $\varphi = 1$  in (24)

We generate i.i.d impulsive Class-A noise samples using the toolbox in [37], while i.i.d Nakagami-m background noise samples by the square root of Gamma distributed r.v..  $MSE\{\hat{\xi}\}$  and  $MSE\{\hat{\eta}\}$  are chosen as the criteria given by  $MSE\{\hat{\xi}\} =$

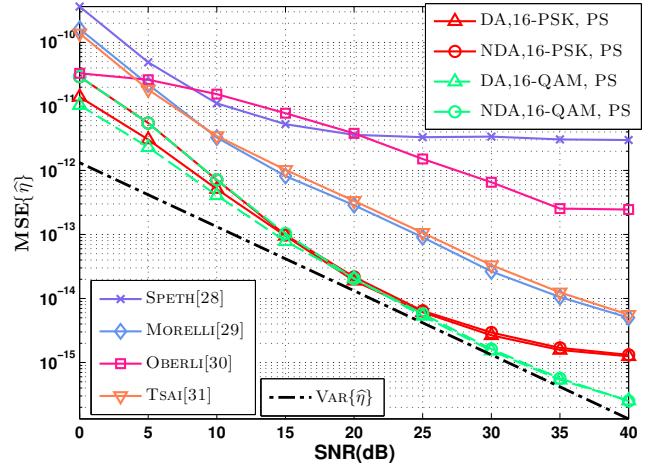


Fig. 4. Comparison of  $MSE\{\hat{\eta}\}$  with PS and other schemes.

$E\{|\hat{\xi} - \xi|^2\}$  and  $MSE\{\hat{\eta}\} = E\{|\hat{\eta} - \eta|^2\}$ . The achievable variances after BE are plotted as  $Var\{\hat{\xi}\}$  and  $Var\{\hat{\eta}\}$  using the first terms of (33) and (34) as benchmarks.

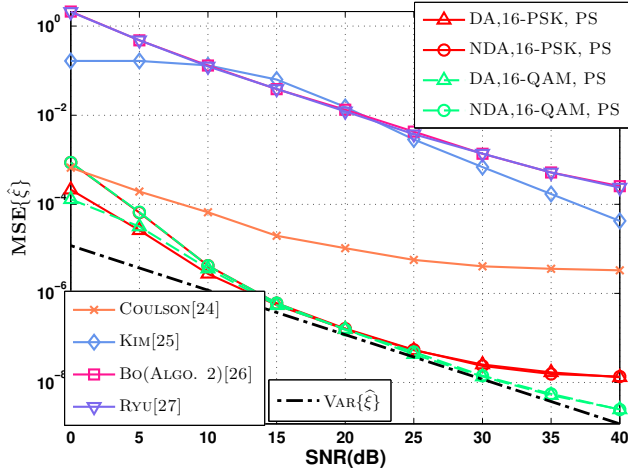
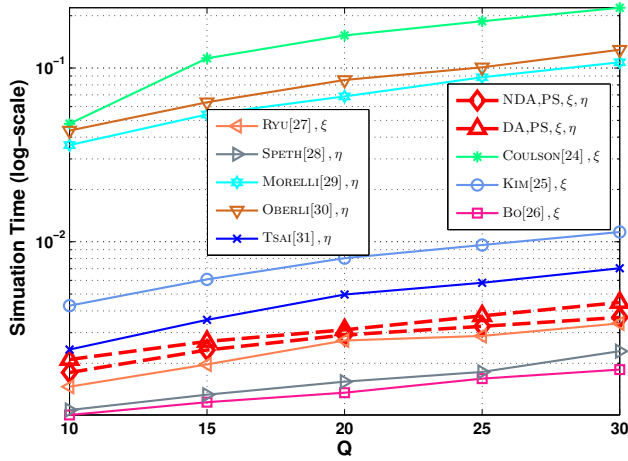
As far as the colored noise, the  $1/f^\beta$  model is used with  $\beta \in \{0.72, 0.337\}$  for generation of correlated noise samples [38]. The condition  $Tr\{\mathbf{B}_n\} = (N-1)\sigma_W^2$  is imposed to keep the total noise power invariant.

For the cyclo-stationary noise, similar to the noise model proposed in [16], the instantaneous variance is set as

$$E\{w(t)\} \triangleq \sigma_W^2(t) = A^2(t) \sin^2\left(\frac{2\pi t}{T_{AC}}\right) \quad (45)$$

where  $T_{AC}$  is the main voltage frequency (60 Hz here).  $A_q$  is chosen according to (46) to make the average energy in the  $q$ -th block  $[(n_0 + (q-1)N)T_{sam}, (n_0 + qN - 1)T_{sam}]$  invariant:

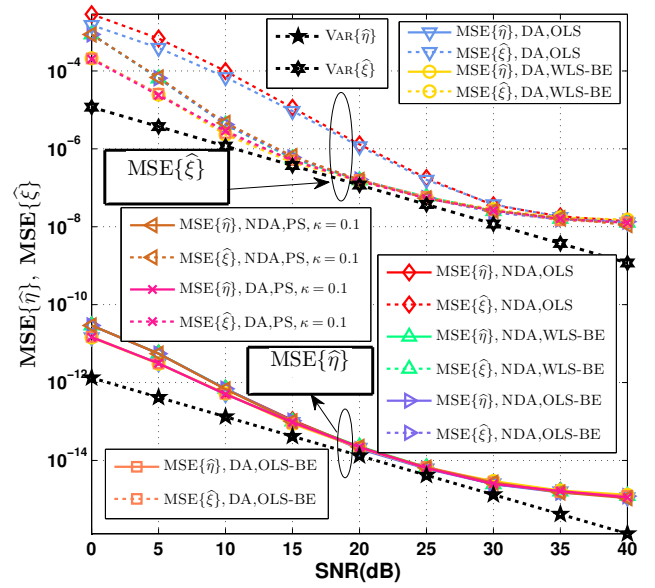
$$A_q = A(t) \Big|_{t=(n_0+(q-1)N)T_{sam}, \dots, (n_0+qN-1)T_{sam}} = \left[ \frac{N\sigma_W^2}{\sum_{n=n_0+(q-1)N}^{n_0+qN-1} \sin^2\left(\frac{2\pi n T_{sam}}{T_{AC}}\right)} \right]^{1/2} \quad (46)$$


 Fig. 5. Comparison of  $\text{MSE}\{\hat{\xi}\}$  with PS and other schemes.

 Fig. 6. Simulation time under different OFDM block number  $Q$ .

#### A. Comparison of $\xi$ and $\eta$ estimation performance using PS with other schemes

Fig. 4 and Fig. 5 demonstrate  $\text{MSE}\{\hat{\eta}\}$  and  $\text{MSE}\{\hat{\xi}\}$  performances of several schemes versus SNR in additive white Gaussian noise (AWGN). All available subcarriers are treated as pilots ( $N_p = N - N_n = 448$ ). Perfect knowledge of CSI is assumed for [24]–[27], [29]–[31] and DA–PS, while unknown for NDA–PS and [28]. For Oberli’s estimator [30],  $\xi = 0$  or otherwise  $\hat{\eta}$  would be biased. Likewise,  $\eta = 0$  for [24]–[27]. RCE is selected from [29], and the *Algorithm 2* out of the four schemes in [26]. Note that, [26] and [27] are originally designed for estimation of both  $\text{Int}\{\xi\}$  and  $\text{Frac}\{\xi\}$  in a multipath channel without knowledge of CSI. However, simulations indicate that CSI should be obtained to guarantee the accuracy of  $\text{Frac}\{\xi\}$ .

PS in (44) is simulated under: (S1) DA, 16-PSK; (S2) NDA, 16-PSK; (S3) DA, 16-QAM; (S4) NDA, 16-QAM. For the performance of  $\hat{\eta}$ , (S1) slightly outperforms (S2) when  $\text{SNR} \leq 20$  dB. Advantage of DA is not significant for moderate SNR since  $\lambda_2(q, k) \approx \text{Const.}$ . The performance gap is more significant in low SNR. For (S3) and (S4), the


 Fig. 7. Different variants of  $\text{MSE}\{\hat{\eta}\}$  and  $\text{MSE}\{\hat{\xi}\}$  estimation including OLS, WLS-BE, OLS-BE, and PS with DA and NDA.

performance gap is enlarged due to the modulation effect of QAM. In the ICI-dominant region, (S3) and (S4) outperform (S1) and (S2).

Similar conclusions can be drawn on  $\hat{\xi}$  except that (S1) slightly outperforms (S3) when  $\text{SNR} \in [4, 15]$  dB. According to the later analysis, the OLS using DA outperforms NDA in low to moderate SNR. However,  $\mu_{\mathcal{A}}^{\theta, \text{SNR}}$ , derived under (A4), departs from the case of QAM. This is more significant for  $\xi$  according to (35) and (36). Therefore, an inaccurate OLS might in reverse compensate the BE inaccuracy and leads to better performance after BE.

As observed in Fig. 4 and Fig. 5, both DA–PS and NDA–PS outperform [24]–[31]. It is even more remarkable considering that the pilot-assisted schemes exploit all usable subcarriers and perfect CSI.

The averaged simulation times of DA–PS, NDA–PS, and [24]–[31] are plotted in Fig. 6. The complexity of the proposed PS schemes are moderate among all the schemes.

#### B. Comparison of the proposed DA/NDA variants

Fig. 7 compares MSE performances of OLS (29), WLS-BE (31), OLS-BE (32), PS (44) using either DA or NDA. For the DA–PS scheme, to simulate imperfect channel estimation and therefore imperfect  $\lambda_2(q, k)$ , a  $N \times 1$  complex Gaussian random vector is introduced which satisfies HSP. This vector, denoted with  $\mathbf{J}_q \triangleq [J_{q,0} \ J_{q,1} \ \cdots \ J_{q,N-1}]^T$ , has zero mean and variance  $\frac{\sum_{k=0}^{N-1} |H_{q,k}|^2}{N}$  for each OFDM block, independent of  $\hat{\mathbf{H}}_q$ . The imperfect CTF, denoted as  $\tilde{\mathbf{H}}_q$ , is expressed by [39]

$$\tilde{\mathbf{H}}_q = \sqrt{1 - \kappa^2} \hat{\mathbf{H}}_q + \kappa \mathbf{J}_q \quad (47)$$

where  $\kappa$  represents the channel estimation accuracy and is set at 0.1 here. Before BE, DA–OLS outperforms NDA–OLS in low SNR. BE is important for  $\hat{\xi}$ , while  $\hat{\eta}$  does not improve after BE mainly due to imperfect phase unwrapping

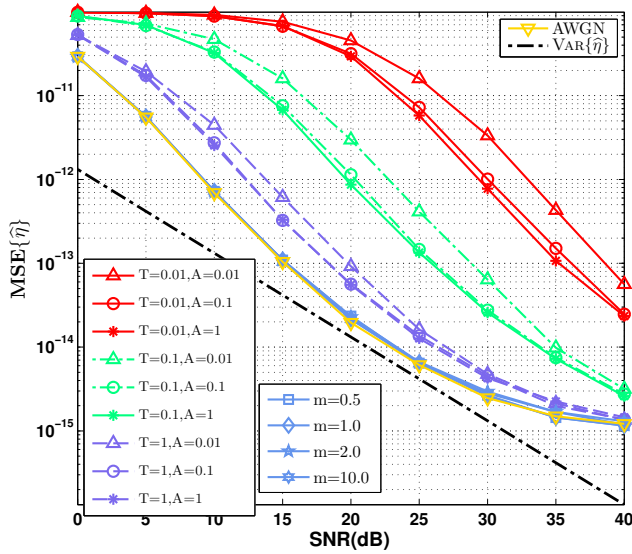


Fig. 8.  $\text{MSE}\{\hat{\eta}\}$  under Class-A and Nakagami-m background noise using NDA and PS.

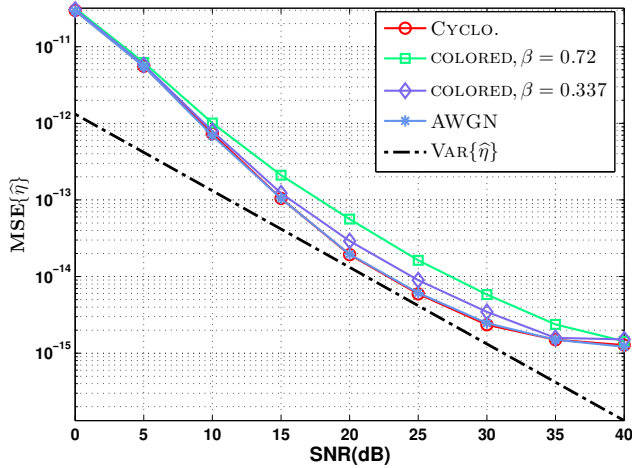


Fig. 9.  $\text{MSE}\{\hat{\eta}\}$  under cyclo-stationarity and temporal correlation using NDA and PS.

not considered in derivations. The performances of WLS-BE, OLS-BE and PS are similar. DA outperforms NDA in all situations for low SNR.  $\kappa$  does not result in any degradation, since  $\lambda_2(q, k)$  is not distorted under (47).

### C. $\text{MSE}\{\hat{\eta}\}$ in Different Noise Models

Fig. 8 shows  $\text{MSE}\{\hat{\eta}\}$  under Class-A impulsiveness and Nakagami-m background noise using NDA-PS; in the former case, SNR is defined as the background SNR as  $\nu\sigma_X^2\sigma_H^2/\sigma_g^2$  to stress the effect of  $T$ .

For the Class-A case, a wide margin exists between  $\text{MSE}\{\hat{\eta}\}$  with different  $T$ . The performance degradation is nearly 7.5 dB when  $T = 1 \Rightarrow T = 0.1$ , and 10 dB when  $T = 0.1 \Rightarrow T = 0.01$ . The gap between AWGN and  $T = 1, A = 1$  is 3 dB.  $A = 0.01$  degrades the performances for  $T = 0.01/0.1$  due to the reduction of Gaussianity shown in Table II, reaching 3.4 dB and 3.3 dB for  $T = 0.1$  and  $T = 0.01$  respectively; a simple remedy is to increase the FFT size  $N$ .

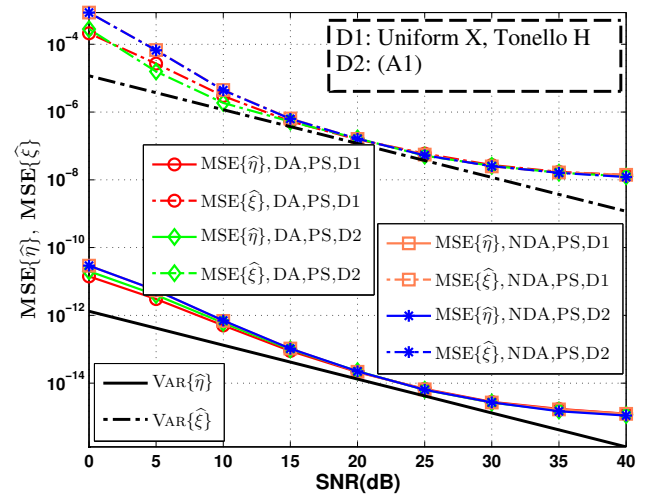


Fig. 10. Comparison of  $\text{MSE}\{\hat{\eta}\}$  and  $\text{MSE}\{\hat{\xi}\}$  under two statistical assumptions.

The losses for  $A = 0.1$  and  $A = 0.01, T = 1$  stem from the increased variance of the numerical result  $\text{Var}\{W\} = \frac{T+1}{T}\sigma_g^2$  given a small  $A$ . In other words, the numerical variance in the Class-A case in Table II converges only in the mean sense for finite size simulations.

As far as the Nakagami-m background noise, performances under different  $m$  are indistinguishable and approach the performance under AWGN, since the noise variance is determined by  $\Omega$  and irrelevant to  $m$ .

Fig. 9 shows  $\text{MSE}\{\hat{\eta}\}$  under cyclo-stationarity and temporal correlation using NDA-PS. The cyclo-stationarity entails no performance loss, since the uniformity of the (anti-)diagonal entries of  $\mathbf{B}_n$  is preserved as mentioned in Section IV. On the other hand, the temporal correlation degrades the performances, which is more pronounced for  $\beta = 0.72$ .

The same conclusions can be made on DA and  $\text{MSE}\{\hat{\xi}\}$  performances which are omitted here due to space limitation.

### D. $\text{MSE}\{\hat{\eta}\}$ and $\text{MSE}\{\hat{\xi}\}$ With Different Statistical Models of Signal and Channel

Fig. 10 compares  $\text{MSE}\{\hat{\eta}\}$  and  $\text{MSE}\{\hat{\xi}\}$  with PS using DA and NDA under

(D1): Uniform distributed  $\mathbf{X}_q$  and  $\dot{\mathbf{H}}_q$  in [6] (Kurtosis: 8.8120)

(D2): Gaussian distributed  $\mathbf{H}_q\mathbf{X}_q$  under (A1) (Kurtosis: 2.9741)

For (D2), samples of  $\mathbf{H}_q\mathbf{X}_q$  are generated directly from  $\mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma_X^2\sigma_H^2\mathbf{I}_N)$ . Interestingly,  $\text{MSE}\{\hat{\eta}\}$  with DA under (D1) slightly outperforms that under (D2), while (D2) slightly outperforms (D1) for  $\text{SNR} \in [2, 15]$  dB in terms of  $\text{MSE}\{\hat{\xi}\}$ . The kurtosis of (D1) is significantly higher than that of (D2), signifying a sharper peak and longer tail distribution. Thus, the probability for a considerably high amplitude  $\mathbf{H}_q\mathbf{X}_q$  under (D1) outnumbers that of (D2), which enhances the profit of  $\lambda_2(q, k)$  using DA. For  $\text{MSE}\{\hat{\xi}\}$ , as stated in V-A, (D2) gains from the reverse compensation. The same performances can be obtained for NDA under (D1) and

(D2). The small performance gap implies the robustness of the proposed technique against violation of (A1) and thus the joint Gaussian assumption of  $\mathbf{R}_q$ .

## VI. CONCLUDING REMARKS

A novel five-step joint timing acquisition technique is proposed to estimate SPO and SCO. An approximation form of the log-likelihood function is derived under the pairwise correlation and Gaussian assumption on the received data in frequency domain. By using the intra-block phase rotations, the proposed technique achieves significant performance improvements over conventional pilot-assisted estimators relying on inter-block phase rotations. Analytical expressions for the variances and biases are derived. Effects of several noise models including non-Gaussianity, cyclo-stationarity, temporal correlation are discussed. The robustness of the proposed technique against violation of the joint Gaussian assumption is verified by simulations.

## VII. ACKNOWLEDGEMENT

This work is supported by National Science and Technology Major Project of China (Grant No. 2011ZX03003-003-03). The authors would like to thank the anonymous reviewers and the editors who help to improve the quality of this paper. The authors also want to express their gratitudes to the members of WiNLAB and BMCLAB of National Central University for their help.

## APPENDIX A

### DERIVATION OF THE APPROXIMATED LOG-LIKELIHOOD FUNCTION IN (19) UNDER (A1) AND (A2)

Using (16) under (A1) and (A2) in Section II,  $\mathbf{R}_q$  is a joint Gaussian vector. Similar to [40], the log-likelihood function is derived as

$$\Lambda(\mathbf{R}'|\boldsymbol{\theta}) = \log \left( \prod_{q=0}^{Q-1} \prod_{k \in \mathcal{I}_2^+} f(R'_{q,k}, R'_{q,N-k}) \right) \quad (\text{A.48})$$

$$f(R'_{q,k}, R'_{q,N-k}) = \frac{1}{\pi^2 \det[\boldsymbol{\Sigma}_G]} \exp(-\mathbf{G}[\boldsymbol{\Sigma}_G]^{-1} \mathbf{G}^H) \quad (\text{A.49})$$

where  $\mathbf{G} \triangleq [R'_{q,k}, R'_{q,N-k}]$ .  $\boldsymbol{\Sigma}_G = \mathbf{E}\{\mathbf{G}^H \mathbf{G}\}$  is given by

$$\boldsymbol{\Sigma}_G = \begin{bmatrix} \sigma_X^2 \sigma_H^2 + \sigma_W^2 & e^{j\Theta_{q,\theta}} \sigma_X^2 \sigma_H^2 + \gamma \sigma_W^2 \\ e^{-j\Theta_{q,\theta}} \sigma_X^2 \sigma_H^2 + \gamma \sigma_W^2 & \sigma_X^2 \sigma_H^2 + \sigma_W^2 \end{bmatrix} \quad (\text{A.50})$$

For convenience,  $\gamma$  is introduced which is 1 for the complete-form log-likelihood function and 0 for the approximated one. The generalized expression for (A.49) is shown in (A.51). For the complete-form log-likelihood function with  $\gamma = 1$ , it is considerably hard to find a closed-form optimal solution to  $\Theta_{q,\theta}$  due to the presence of  $[1 - \cos(\Theta_{q,\theta})]$ . Thus,  $\gamma$  is set at 0 which gives (19). To fill the gap between the complete and the approximated form of the log-likelihood function,  $\sigma_X^2$  and  $\sigma_H^2$  are replaced by  $|X_{q,k}|^2$  and  $|H_{q,k}|^2$  respectively to take advantage of the modulation effect and frequency selectivity, which gives  $\lambda_2(q, k)$  in (21).

## APPENDIX B

### DERIVATION OF VARIANCES AND BIASES IN (33), (34), (35), (36) UNDER (A2), (A3), AND (A4)

#### A. Bias Performance

For DA, subtracting the left hand side of (22) by the true value of  $\Theta_{q,\theta}$  yields

$$\Delta\Theta_{q,\theta} = \widehat{\Theta}_{q,\theta}^{\text{DA}} - \Theta_{q,\theta} = \arg \left( \sum_{k \in \mathcal{I}_2^+} \frac{\lambda_1(q, k) \lambda_2(q, k)}{e^{j\Theta_{q,\theta}}} \right) \quad (\text{B.52})$$

Suppose that  $\widehat{\Theta}_{q,\theta}^{\text{DA}}$  is in vicinity of  $\Theta_{q,\theta}$ ,  $\tan(\Delta\Theta_{q,\theta})$  can be approximated as

$$\tan(\Delta\Theta_{q,\theta}) = \frac{\Im[\Delta\Theta_{q,\theta}]}{\Re[\Delta\Theta_{q,\theta}]} \approx \Delta\Theta_{q,\theta} \quad (\text{B.53})$$

since  $\Delta\Theta_{q,\theta}$  is sufficiently small. Expectation of  $\Delta\Theta_{q,\theta}$  in (B.53) is approximated by

$$\mathbf{E}\{\Delta\Theta_{q,\theta}\} \approx \mathbf{E}\left\{ \frac{\Im[\Delta\Theta_{q,\theta}]}{\Re[\Delta\Theta_{q,\theta}]} \right\} \approx \frac{\mathbf{E}\{\Im[\Delta\Theta_{q,\theta}]\}}{\mathbf{E}\{\Re[\Delta\Theta_{q,\theta}]\}} \quad (\text{B.54})$$

if  $\mathbf{E}\{\Re[\Delta\Theta_{q,\theta}]\} \gg \sqrt{\text{Var}\{\Re[\Delta\Theta_{q,\theta}]\}}$ . Under (A2), write

$$\mathbf{E}\{\Im[\Delta\Theta_{q,\theta}]\} = - \sum_{k \in \mathcal{I}_2^+} \sin(\Theta_{q,\theta}) \lambda_2(q, k) \quad (\text{B.55})$$

$$\mathbf{E}\{\Re[\Delta\Theta_{q,\theta}]\} = \sum_{k \in \mathcal{I}_2^+} [\cos(\Theta_{q,\theta}) + \text{SNR}(q, k)] \lambda_2(q, k) \quad (\text{B.56})$$

Under (A3) and (A4),  $\mathbf{E}\{\Delta\Theta_{q,\theta}\}$  takes the form

$$\mathbf{E}\{\Delta\Theta_{q,\theta}\} \approx - \frac{\sin(\Theta_{q,\theta})}{\cos(\Theta_{q,\theta}) + \text{SNR}} \quad (\text{B.57})$$

where  $\text{SNR} = \text{SNR}(q, k)$  and  $\lambda_2(q, k) = (2 + \text{SNR}^{-1}) = \text{Const.}$ . Assuming perfect phase unwrapping, (29) gives  $\text{Bias}\{\hat{\boldsymbol{\theta}}_{\text{DA}}\} = \text{Bias}\{\mathbf{V}\}$ . The  $2 \times 1$  bias vector  $\text{Bias}\{\mathbf{V}\}$  is associated with the  $Q \times 1$  bias vector  $\text{Bias}\{\mathbf{V}\}$  by

$$\text{Bias}\{\mathbf{V}\} = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \text{Bias}\{\mathbf{V}\} \quad (\text{B.58})$$

where  $\text{Bias}\{\mathbf{V}\} = \begin{bmatrix} -\frac{\sin(\Theta_{0,\theta})}{\cos(\Theta_{0,\theta}) + \text{SNR}} \\ \frac{\sin(\Theta_{1,\theta})}{\cos(\Theta_{1,\theta}) + \text{SNR}} \cdots -\frac{\sin(\Theta_{Q-1,\theta})}{\cos(\Theta_{Q-1,\theta}) + \text{SNR}} \end{bmatrix}^T$ . Standard calculations lead to (35) and (36).

#### B. Variance Performance

$\text{Var}\{\Delta\Theta_{q,\theta}\}$  can be formulated into

$$\text{Var}\{\Delta\Theta_{q,\theta}\} \approx \frac{\text{Var}\{\Im[\Delta\Theta_{q,\theta}]\}}{(\mathbf{E}\{\Re[\Delta\Theta_{q,\theta}]\})^2} \quad (\text{B.59})$$

under  $\mathbf{E}\{\Re[\Delta\Theta_{q,\theta}]\} \gg \sqrt{\text{Var}\{\Re[\Delta\Theta_{q,\theta}]\}}$ . Again, assuming (A2),  $(\mathbf{E}\{\Re[\Delta\Theta_{q,\theta}]\})^2$  can be written into (B.60).

Note that,  $\text{Var}\{\Im[\Delta\Theta_{q,\theta}]\} = \mathbf{E}\{(\Im[\Delta\Theta_{q,\theta}])^2\} - (\mathbf{E}\{\Im[\Delta\Theta_{q,\theta}]\})^2$  where

$$(\mathbf{E}\{\Im[\Delta\Theta_{q,\theta}]\})^2 = \sin^2(\Theta_{q,\theta}) \sigma_W^4 \left( \sum_{k \in \mathcal{I}_2^+} \lambda_2(q, k) \right)^2 \quad (\text{B.65})$$

$$f(R'_{q,k}, R'_{q,N-k}) = \frac{\exp\left(\frac{-(\sigma_X^2\sigma_H^2 + \sigma_W^2)(|R_{q,k}|^2 + |R_{q,N-k}|^2) + 2\Re\{R_{q,k}R_{q,N-k}e^{-j\Theta_{q,\theta}}\sigma_X^2\sigma_H^2\}}{2\sigma_X^2\sigma_H^2\sigma_W^2(1-\gamma\cos(\Theta_{q,\theta})) + \sigma_W^4(1-\gamma^2)}\right)}{\pi^2[2\sigma_X^2\sigma_H^2\sigma_W^2(1-\gamma\cos(\Theta_{q,\theta})) + \sigma_W^4(1-\gamma^2)]} \quad (\text{A.51})$$

$$(\mathbb{E}\{\Re[\Delta\Theta_{q,\theta}]\})^2 = \left[ \sum_{k \in \mathcal{I}_2^+} (|X_{q,k}|^2 |H_{q,k}|^2 + \sigma_W^2 \cos(\Theta_{q,\theta})) \lambda_2(q, k) \right]^2 \quad (\text{B.60})$$

$$\begin{aligned} (\Im[\Delta\Theta_{q,\theta}])^2 = & \sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} \left( \underbrace{|W_{q,k_1}|^2 |W_{q,k_2}|^2 \sin^2(\Theta_{q,\theta})}_{Y_1} - \underbrace{|W_{q,k_1}|^2 \sin(\Theta_{q,\theta}) \Im\{\lambda_3(q, k_2)\}}_{Y_2} \right. \\ & \left. - \underbrace{|W_{q,k_2}|^2 \sin(\Theta_{q,\theta}) \Im\{\lambda_3(q, k_1)\}}_{Y_3} + \underbrace{\Im\{\lambda_3(q, k_1)\} \Im\{\lambda_3(q, k_2)\}}_{Y_4} \right) \times \underbrace{\lambda_2(q, k_1) \lambda_2(q, k_2)}_{Y_5} \quad (\text{B.61}) \end{aligned}$$

$$\lambda_3(q, k) = e^{j\Theta_{q,k-N,\xi,\eta}} X_{q,k} H_{q,k} W_{q,N-k} + e^{j\Theta_{q,-k,\xi,\eta}} X_{q,N-k} H_{q,N-k} W_{q,k} \quad (\text{B.62})$$

$$\mathbb{E}\left\{ \sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} Y_1 \times Y_5 \right\} = \sin^2(\Theta_{q,\theta}) \sigma_W^4 \sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} \lambda_2(q, k_1) \lambda_2(q, k_2) \quad (\text{B.63})$$

$$\mathbb{E}\left\{ \sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} Y_4 \times Y_5 \right\} \approx \sigma_W^2 (1 - \cos(\Theta_{q,\theta})) \sum_{k \in \mathcal{I}_2^+} |X_{q,k}|^2 |H_{q,k}|^2 \lambda_2^2(q, k) \quad (\text{B.64})$$

Expanding  $(\Im[\Delta\Theta_{q,\theta}])^2$  into the form shown as (B.61) with  $\lambda_3(q, k)$  shown as (B.62).  $\mathbb{E}\{\sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} Y_1 \times Y_5\}$  and  $\mathbb{E}\{\sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} Y_4 \times Y_5\}$  can be computed as (B.63) and (B.64), while  $\mathbb{E}\{\sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} Y_2 \times Y_5\}$  and  $\mathbb{E}\{\sum_{k_1 \in \mathcal{I}_2^+} \sum_{k_2 \in \mathcal{I}_2^+} Y_3 \times Y_5\}$  can be neglected.  $\text{Var}\{\Im[\Delta\Theta_{q,\theta}]\}$  under (A3) and (A4) is expressed by

$$\text{Var}\{\Im[\Delta\Theta_{q,\theta}]\} \approx \Xi N \frac{\text{SNR} (1 - \cos(\Theta_{q,\theta}))}{(2 + \text{SNR}^{-1})^2} \quad (\text{B.66})$$

Substituting  $\text{Var}\{\Im[\Delta\Theta_{q,\theta}]\}$  and  $(\mathbb{E}\{\Re[\Delta\Theta_{q,\theta}]\})^2$  back into (B.59) yields

$$\text{Var}\{\Delta\Theta_{q,\theta}\} \approx \frac{1 - \cos(\Theta_{q,\theta})}{\Xi N \times \text{SNR}} \quad (\text{B.67})$$

Therefore,

$$\text{Var}\{\widehat{\boldsymbol{\theta}}_{\text{DA}}\} = \text{Var}\{\mathbf{V}\} = \text{Var}\{(\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \mathbf{V}\} \quad (\text{B.68})$$

Lengthy calculations lead to (33) and (34). Under (A3) and (A4), the DA estimator reduces to NDA.

## REFERENCES

- [1] S. Galli and O. Logvinov, "Recent developments in the standardization of power line communications within the IEEE," *IEEE Commun. Mag.*, vol. 46, no. 7, pp. 64–71, july 2008.
- [2] H. Meng, Y. Guan, and S. Chen, "Modeling and analysis of noise effects on broadband power-line communications," *IEEE Trans. Power Delivery*, vol. 20, no. 2, pp. 630–637, april 2005.
- [3] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 998–1027, june 2011.
- [4] S. Galli, "Power line communications," in *IEEE Communication Theory Workshop, Cancun, Mexico*, may 2010.
- [5] M. Zimmermann and K. Dostert, "A multipath model for the powerline channel," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 553–559, apr 2002.
- [6] A. Tonello, F. Versolatto, B. Bejar, and S. Zazo, "A fitting algorithm for random modeling the PLC channel," *IEEE Trans. Power Delivery*, vol. 27, no. 3, pp. 1477–1484, july 2012.
- [7] F. Canete, J. Cortés, L. Díez, and J. Entrambasaguas, "A channel model proposal for indoor power line communications," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 166–174, december 2011.
- [8] S. Barmada, A. Musolino, M. Raugi, R. Rizzo, and M. Tucci, "A wavelet based method for the analysis of impulsive noise due to switch commutations in power line communication (PLC) systems," *Smart Grid, IEEE Transactions on*, vol. 2, no. 1, pp. 92–101, march 2011.
- [9] Z. Tao, Y. Xiaoxian, Z. Baohui, N. Xu, F. Xiaoqun, and L. Changxin, "Statistical analysis and modeling of noise on 10-kV Medium-Voltage power lines," *IEEE Trans. Power Delivery*, vol. 22, no. 3, pp. 1433–1439, Jul. 2007.
- [10] Y. Kim, S. Choi, and H. Oh, "Closed-Form expression of Nakagami-Like background noise in Power-Line channel," *IEEE Trans. Power Delivery*, vol. 23, no. 3, pp. 1410–1412, Jul. 2008.
- [11] Y. S. Kim, Y. H. Kim, H. M. Oh, and S. Choi, "BER performance of binary transmitted signal for power line communication under Nakagami-like background noise," in *ENERGY 2011, The First International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies*, May 2011, pp. 126–129.
- [12] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods and results for Class A and Class B noise models," *IEEE Trans. Inform. Theory*, vol. 45, no. 4, pp. 1129–1149, may 1999.
- [13] L. Berry, "Understanding Middleton's canonical formula for Class A noise," *IEEE Trans. Electromagn. Compat.*, vol. EMC-23, no. 4, pp. 337–344, Nov. 1981.
- [14] R. Haring and A. Han Vinck, "Performance bounds for optimum and suboptimum reception under Class-A impulsive noise," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1130–1136, jul 2002.
- [15] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Electro-*

- magn. Compat.*, vol. 44, no. 1, pp. 249–258, feb 2002.
- [16] M. Katayama, T. Yamazato, and H. Okada, “A mathematical model of noise in narrowband power line communication systems,” *IEEE J. Select. Areas Commun.*, vol. 24, no. 7, pp. 1267–1276, july 2006.
- [17] R. Mo, S. W. Oh, Y. Zeng, and S. Lie, “Channel estimation for power line OFDM systems with unknown colored noise exploiting real signal transmission,” in *Power Line Communications and Its Applications, 2009. ISPLC 2009. IEEE International Symposium on*, Apr. 2009, pp. 149–153.
- [18] T. Pollet and M. Peeters, “Synchronization with DMT modulation,” *IEEE Commun. Mag.*, vol. 37, no. 4, pp. 80–86, apr 1999.
- [19] Y.-H. Kim and J.-H. Lee, “Comparison of passband and baseband transmission schemes for power-line communication OFDM systems,” *Power Delivery, IEEE Transactions on*, vol. 26, no. 4, pp. 2466–2475, oct. 2011.
- [20] T.-D. Chiueh and P.-Y. Tsai, *OFDM Baseband Receiver Design for Wireless Communications*. John Wiley and Sons (Asia) Pte Ltd, 2007.
- [21] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM wireless communications with MATLAB*. John Wiley and Sons (Asia) Pte Ltd, 2010.
- [22] J. Cortés, L. Díez, E. Martos, F. Canete, and J. Entrambasaguas, “Analysis of timing recovery for DMT systems over indoor power-line channels,” in *Global Telecommunications Conference, 2006. GLOBE-COM '06. IEEE*, 27 2006-dec. 1 2006, pp. 1–6.
- [23] B. Yang, K. Letaief, R. Cheng, and Z. Cao, “Timing recovery for OFDM transmission,” *IEEE J. Select. Areas Commun.*, vol. 18, no. 11, pp. 2278–2291, nov. 2000.
- [24] A. Coulson, “Maximum likelihood synchronization for OFDM using a pilot symbol: algorithms,” *IEEE J. Select. Areas Commun.*, vol. 19, no. 12, pp. 2486–2494, dec 2001.
- [25] D. K. Kim, S. H. Do, H. B. Cho, H. J. Chul, and K. B. Kim, “A new joint algorithm of symbol timing recovery and sampling clock adjustment for OFDM systems,” *IEEE Trans. Consumer Electron.*, vol. 44, no. 3, pp. 1142–1149, aug 1998.
- [26] B. Ai, Z. xing Yang, C. yong Pan, J. hua Ge, Y. Wang, and Z. Lu, “On the synchronization techniques for wireless OFDM systems,” *IEEE Trans. Broadcast.*, vol. 52, no. 2, pp. 236–244, june 2006.
- [27] Y.-J. Ryu and D.-S. Han, “Timing phase estimator overcoming Rayleigh fading for OFDM systems,” *IEEE Trans. Consumer Electron.*, vol. 47, no. 3, pp. 370–377, aug 2001.
- [28] M. Speth, S. Fechtel, G. Fock, and H. Meyr, “Optimum receiver design for OFDM-based broadband transmission II: A case study,” *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 571–578, apr 2001.
- [29] M. Morelli and M. Moretti, “Fine carrier and sampling frequency synchronization in OFDM systems,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1514–1524, april 2010.
- [30] C. Oberli, “ML-based tracking algorithms for MIMO-OFDM,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2630–2639, july 2007.
- [31] P.-Y. Tsai, H.-Y. Kang, and T.-D. Chiueh, “Joint weighted least-squares estimation of carrier-frequency offset and timing offset for OFDM systems over multipath fading channels,” *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 211–223, jan. 2005.
- [32] M. Speth, S. Fechtel, G. Fock, and H. Meyr, “Optimum receiver design for wireless broad-band systems using OFDM—Part I,” *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1668–1677, nov 1999.
- [33] S. Aghajeri, M. R. Gholami, and H. Nikopoor, “Blind sampling clock adjustment in OFDM power line communication,” in *9th International Symposium on Signal Processing and Its Applications, 2007. ISSPA 2007*. IEEE, Feb. 2007, pp. 1–4.
- [34] N. Beaulieu and C. Cheng, “Efficient Nakagami-m fading channel simulation,” *IEEE Trans. Veh. Commun.*, vol. 54, no. 2, pp. 413–424, march 2005.
- [35] J. Cheng and N. Beaulieu, “Maximum-likelihood based estimation of the Nakagami m parameter,” *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 101–103, march 2001.
- [36] R. B. D’Agostino and M. A. Stephens, *Goodness-of-fit-techniques*. Dekker, 1986.
- [37] K. Gulati, M. Nassar, A. Chopra, N. Okafor, M. DeYoung, A. Sujeeth, N. Aghasadeghi, and B. Evans, “UT Austin interference modeling and mitigation toolbox,” Department of Electrical and Computer Engineering The University of Texas at Austin, Austin, Texas USA, Tech. Rep., Mar. 2012.
- [38] T. Esmailian, F. R. Kschischang, and P. Glenn Gulak, “In-building power lines as high-speed communication channels: channel characterization and a test channel ensemble,” *International Journal of Communication Systems*, vol. 16, no. 5, pp. 381–400, 2003.
- [39] K. S. Ahn, R. Heath, and H. K. Baik, “Shannon capacity and symbol error rate of space-time block codes in MIMO Rayleigh channels with channel estimation error,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 324–333, jan. 2008.
- [40] J. van de Beek, M. Sandell, and P. Borjesson, “ML estimation of time and frequency offset in OFDM systems,” *IEEE Trans. Signal Processing*, vol. 45, no. 7, pp. 1800–1805, jul 1997.